Cost and Performance Trade-offs in Reconfiguration Strategies for WDM Networks

Cost and Performance Trade-offs in Reconfiguration Strategies for WDM Networks

Andrea Bianco #, Jorge Finochietto *, Chiara Piglione †

#*Dip. di Elettronica, Politecnico di Torino*
Corso Duca degli Abruzzi 24, 10129 Torino, Italy
andrea.bianco@polito.it

**CONICET - Universidad Nacional de Cordoba*
Cordoba, Argentina,
jorge.finochietto@ieee.org

**DCBU Modelling and Simulation Team, CISCO Systems*
San José, USA
cpiglion@cisco.com

***Abstract*— In this paper we study the problem of receiver allocation in WDM single-hop optical networks under dynamically changing traffic. Instead of focusing only on performance optimization, we wish to take into account both network performance and reconfiguration cost reduction. We first formalize the receiver allocation problem as a multi-criteria ILP. Then, we define Pareto-optimal solutions to solve the problem. The number and characteristics of Pareto-optimal solutions is studied to assess the problem significance. Finally, we propose different reconfiguration schemes that may guide the choice of network administrators among efficient solutions for the re-allocation problem.**

I. Introduction

Single-hop WDM optical networks operating in packet mode are considered a promising choice for future Metros. The single-hop approach avoids complex switching in the optical domain and thus permits a cost-effective balance of optics and electronics. In these networks, nodes are equipped with few (typically one) transceivers, and each transceiver operates at the data rate of a single WDM channel. Tunability at transceivers is required to exploit the fiber bandwidth by temporally (on a packet by packet basis) allocating all-optical single-hop bandwidth between nodes in all available WDM channels, according to a dynamic TDMA access scheme on each WDM channel.

Due to the cost of tunability at transceivers, media access protocols that require packet-by-packet tunability only either at the transmitter or at the receiver have been studied, to save the cost of the still quite expensive tunable devices. Usually nodes are equipped with a *fastly* tunable transmitter and a fixed receiver, permanently tuned to a WDM channel. When a node needs to send a packet, it simply tunes its transmitter to the receiver's destination wavelength. This implies that transmitter tuning times must be negligible with respect to the packet duration to obtain a good efficiency.

If the number of nodes N is larger than the number of WDM channels W, a decision problem arises concerning the allocation of node receivers to WDM channels, with the aim of balancing the traffic among the available WDM channels. Indeed, let us suppose that the traffic pattern is described via a demand matrix $\mathbf{T}$ of size $N \times N$, describing the amount of traffic each node is willing to send to other nodes, is known either by estimation or measurement. Node receivers must be assigned to one among the W available channels. Obviously, traffic demands must be taken into account to avoid channel overloading and/or to balance load on different channels. Receivers assigned to the same WDM channel will share the bandwidth according to a dynamic TDMA scheme. If fixed receivers are considered, any allocation is fixed and cannot be updated in response to changes in the traffic pattern, which are typical in Metros. Therefore, it may be worthwhile to re-allocate, i.e. tune to different wavelengths, receivers to keep the network in an optimal operation point. One elegant way of achieving this result is to introduce *slow* (hence cheap) tunability in receivers. This tunability does not need to be *fast*, since it must not track packet-by-packet variations, but longer-term variations of the traffic pattern.

A typical solution to the receiver allocation problem aims at optimizing performance by balancing the load on all available channels or by determining an allocation that does not overload any channel [1]. This problem, a particular case of a logical topology design problem, can be shown to be equivalent to the problem of scheduling jobs on parallel machines [2], a single criterion ILP (Integer Linear Programming) problem. However, this approach neglects transceiver reconfiguration cost. Indeed, slow tuning latencies on receivers introduce a period of blackout, since all nodes must refrain from transmitting to destinations under reconfiguration. Under lightly loaded conditions, blackout periods affect mainly transmission delays, while under high load conditions packet losses can occur.

Normally, it is assumed that blackout periods are much shorter with respect to traffic fluctuations, so that this phenomenon can be neglected when determining a new receiver allocation via reconfiguration. However, in some situations the receiver reconfiguration process may require longer blackout

978-1-4244-1845-9/08/$25.00

periods (e.g. in linear topologies a RTT may be necessary to ensure that all nodes are correctly allocated to the proper WDM channel). This implies that the performance penalty created by the blackout period may become significant. Moreover, several different re-allocation schemes may provide similarly acceptable performance (i.e., no overloaded channel) with different reconfiguration costs (e.g., numbers of receivers to be tuned). Finally, in some cases the network administrator may be willing to control not only performance but also disruption cost, that may be difficult to quantify since they influence user satisfaction. Indeed, during node receiver re-allocation to a different wavelength channel, nodes are not active on the data path, creating service interruption. For all these reasons, it may be interesting to study a scenario when the traffic matrix changes over time and receiver allocation schemes take into account both reconfiguration costs and performance, leading to a multi-criteria ILP problem.

Multi-criteria problems can be reduced to single-criterion problems by linearization. However, this requires first the capability to define proper coefficients weighting the single criterion; second, criteria should be commensurable. This is not our case, since reconfiguration cost and performance cannot be directly compared. As such, a single optimal solution cannot be easily defined. In this context, a set of solutions can solve efficiently the receiver allocation problem and it may be difficult to automatically determine a ranking criteria among those solutions. This set of solutions, represented as vectors of criteria, is called Pareto-optimal set. All the Pareto-optimal solutions of interests must be efficient, i.e., no other solution vector can decrease some criterion without causing a simultaneous increase in at least one of the other criterion.

In this paper, first, we analyze how many Pareto-optimal solutions are available in a typical network reconfiguration scenario, so as to assess the problem significance. Then, we analyze the characteristics of the Pareto-optimal solutions. Finally, we provide algorithms, named reconfiguration strategies, that may be useful to guide the choice of network administrators among efficient solutions, since no automatic procedure to determine a ranking among these solutions can be easily defined.

II. Problem formalization

Let us formalize the receiver allocation problem as an ILP (Integer Linear Programming) model. We refer to a network with N nodes, W channels, with $N > W$; each node is equipped with a single fixed transmitter and a single tunable receiver.

Introduce the set of control variables x_{ik}, where:

$$x_{ik} = \begin{cases} 1 & \text{iff node } i \text{ receives on wavelength } k \\ 0 & \text{otherwise} \end{cases}$$

It is straightforward to notice that the solution of the receiver allocation problem depends on source-destination node bandwidth requirements. Denote by $\mathbf{T} = [t_{ij}]$ the traffic demand matrix, where t_{ij} is the amount of traffic to be transmitted from node *i* to node *j*. Since the problem focuses on receivers allocation, we are interested on the receiver aggregated traffic t_j :

$$t_j = \sum_{i=1}^{N} t_{ij} \qquad \forall i,\ 1 \le i \le N \tag{1}$$

We first consider the single-criterion problem, where the only objective is load minimization, and later extend it to the multi-criteria problem, where also reconfiguration costs are taken into account.

A. The single-criterion problem

In the *single-criterion* case, we first assume that the traffic matrix $\mathbf{T}$ does not change over time. Thus, reconfiguration costs are neglected, and the only allocation criteria is to minimize $\mathcal{L}_{\max}$, i.e. the load on the most loaded wavelength, being $\mathcal{L}_{\max} = \max_k \sum_{i=1}^{N} t_i x_{ik}$.

Thus, the problem becomes:

$$\text{Minimize } \mathcal{L}_{\max} \tag{2}$$

subject to the following constraints:

$$\mathcal{L}_{\max} \ge \sum_{i=1}^{N} t_i x_{ik} \qquad \forall k,\ 1 \le k \le W \tag{3}$$

$$\sum_{k=1}^{W} x_{ik} = 1 \qquad \forall i,\ 1 \le i \le N \tag{4}$$

Eq. (3) guarantees that no wavelength has a load larger than $\mathcal{L}_{\max}$, while Eq. (4) ensures that each receiver is allocated only to one wavelength, since nodes are equipped with a single receiver.

Note that slightly different optimization problems can be defined by modifying the minimization criterion. We focus on the minimization of the load on the most loaded wavelength. This choice guarantees that there are no overloaded wavelength channels in the network, if such an allocation exists. Moreover, by minimizing the load on the most loaded channel, the resulting allocation is relatively robust to node-to-node traffic variations with respect to the nominal values contained in the traffic demand matrix.

This problem is equivalent to scheduling jobs on identical parallel machines. Receiver's aggregated traffic t_i represents job's duration, and wavelengths represent machines. Although it falls in the class of NP-hard problems [3], an approximation algorithm known as *Largest Processing Time* (LPT) [4] limits the distance from the optimal solution $\mathcal{L}^*_{\max}$ such that

$$\mathcal{L}^{\text{LPT}}_{\max} \le \left(\frac{4}{3} - \frac{1}{3W}\right) \mathcal{L}^*_{\max} \tag{5}$$

The LPT algorithm is simple, and runs through the following steps:

1 Sort nodes by decreasing t_i, $\forall i = 1...N$.
2 Allocate largest (unallocated) t_i to least loaded channel.
3 If unallocated receivers exist, go to 2.

If the traffic matrix $\mathbf{T}$ changes over time, a new LPT algorithm may be run for each traffic variation to determine the new optimal receiver allocation.

B. The Multi-criteria problem

In the *multi-criteria* case, we assume that the traffic matrix changes over time and that the receiver allocation tracks these variations, trying to optimize bandwidth usage while minimizing reconfiguration cost.

To account for reconfiguration costs, we define c_{ik} as the cost of re-allocating node i to wavelength k. Note that if a receiver has to be tuned to a wavelength different from the one where it is currently tuned, all nodes willing to transmit to this destination must refrain from transmissions for a period of time equal at least to the tuning latency. Once we have defined the reconfiguration costs, we can modify the single-criterion ILP model to take into account the costs, expressed, for example, by $\mathcal{C} = \sum_{i=1}^{N} \sum_{k=1}^{W} c_{ik} x_{ik}$; therefore, the receiver allocation problem becomes a multi-criteria problem.

$$\text{Minimize } [\mathcal{L}_{\max},\, \mathcal{C}] \tag{6}$$

subject to constraints (3) and (4).

Note that the above expression of $\mathcal{C}$ is simply a possible example of a cost function. Indeed, adding costs may have sense if considering disruption cost, i.e. the inability to exploit network resources during network reconfiguration, but it may be meaningless in different scenarios. The choice of the best cost function should be made by network administrators, taking into account the specific needs of the network under control.

Unfortunately, multi-criteria problems are hard to be dealt with; typically, the trade-off is among the so-called efficient, or Pareto-optima, solutions. Since the objective function is no more a scalar but a vector, the optimality concept used in scalar optimization is replaced by a new one, called Pareto-optimum. A vector representing a possible solution to the problem is said to be Pareto-optimal if no other solution vector can decrease some criterion without causing a simultaneous increase in at least another criterion. In other words, given a feasible solution, if we cannot find a lower $\mathcal{L}_{\max}$ without increasing the cost $\mathcal{C}$, or vice-versa, this solution is Pareto-optimum, since, to lower $\mathcal{L}_{\max}$, we have to accept to pay an increase in reconfiguration cost.

The solution to the optimization problem is not a single vector, but a set of Pareto optimal vectors. This fact implies that although we know all Pareto-optima, we still have to decide which one represents the best solution. In the following sections, we first discuss the problem of finding Pareto-optimal solutions, and, later, we propose an algorithm that may help in the decision process of selecting the final allocation.

III. Receiver Allocation Algorithms

The most common way to treat a multi-criteria problem is using a scalar optimization approach. A good example of this method is a linear combination of the different criteria, maybe using different coefficients, into one objective function. However, this approach does not guarantee the generation of the whole set of Pareto-optimal solutions, besides being sometimes difficult to apply when criteria are not commensurable as in our scenario. Due to its simplicity, it is useful when the cost of generating a single Pareto optimum solution may become so high that the designer can afford only a few Pareto-optimal solutions. In contrast, when the cost of generating Pareto-optima is low, we can generate the complete set and then decide which solution to choose.

A. Pareto Search

A way of searching Pareto-optimal solutions is to use the so-called ε-constraint method [5]. Let $\bar{\sigma}$ be our current allocation where no reconfiguration is done (i.e., $\mathcal{C} = 0$) and whose maximum load is $\mathcal{L}_{\max}(\bar{\sigma})$, and let $L_{max}(SC)$ be the solution to the *single-criterion* case, where the aggregate load is minimized. Notice that for any Pareto-optimal allocation σ, the value of C belongs to the interval $[\mathcal{C}(\bar{\sigma}), \mathcal{C}(SC)]$. The method chooses one criterion as a scalar objective to be minimized and transform the other one into a constraint. Thus, we minimize the cost for different C values by applying the following loop:

Pareto-Search Algorithm
Set Pareto List $PL = \{\bar{\sigma}\}$, budget $B = \varepsilon$ and $\theta = \bar{\sigma}$;
while $\mathcal{L}_{\max}(\theta) >= \mathcal{L}_{\max}(SC)$
1 Solve $P(B)\ :\ \min \mathcal{L}_{\max}(\sigma)$ when $\mathcal{C}(\sigma) \leq B$
2 Let $\theta = arg\ \min_{\pi \in PL}\{\mathcal{L}_{\max}(\pi)\}$
3 If $\mathcal{L}_{\max}(\sigma) < \mathcal{L}_{\max}(\theta)$ THEN add σ to PL
4 Increase budget: $B = B + \varepsilon$
end while

The algorithm is capable of searching all Pareto-optimal solutions when ε is small enough. Notice that all Pareto solutions are added to the Pareto List at Step 3 and represent possible solutions to our problem. Besides, Pareto-optimal solutions having the same cost C and $\mathcal{L}_{\max}$ are equivalent, since no other criterion is specified; thus, the algorithm searches only non-equivalent Pareto-optimal solutions.

To apply the Pareto-Search algorithm, we need an efficient algorithm to solve $P(B)$. The branch and bound technique provides an exact algorithm to reach the optimal solution; however, when the number of variables (equal to NW) is large, a heuristic could represent a good trade-off between complexity and performance. We rely on branch and bound in the remainder of the paper to determine Pareto-optimal solutions, since we analyze performance for relatively small networks to discuss the problem significance. Heuristics to determine Pareto-optimal solutions should clearly be adopted when dealing with real-size networks.

In the following subsection, we discuss criteria for selecting the best solution among Pareto-optima ones.

B. Reconfiguration Strategies

We propose four reconfiguration strategies that may be used to select one among the Pareto-optimal solutions determined in the Pareto search procedure in our two-criteria minimization problem.

1) Minimum Cost: This solution is trivial and represents the case in which, even if a reconfiguration process may improve network performance, the current allocation state is preferred to avoid any reconfiguration. This strategy may have sense when the cost criterion is considered dominant: for example when a certain Service Level Agreement (SLA) or availability constraint must be satisfied.

2) Minimum Overload: This strategy follows the opposite approach to the previous one. We aim at the best solution in terms of load balancing among channels, trying to pay the minimum cost to obtain this solution. If this strategy is always preferred, instead of generating all optimal solutions, we can generate only this Pareto-optimal solution by decoupling the problem. Indeed, we could first solve the *single-criterion* case for current traffic conditions, and then find the least expensive reconfiguration to move from the current allocation to the best one.

3) Best Ratio: Define $\mathcal{L}_{\max}(\bar{\sigma})$ the maximum load of the current allocation, $\mathcal{L}_{\max}(Pareto)$ and $\mathcal{C}(Pareto)$, respectively, the maximum load and the cost of the Pareto-optimal solution under examination. We select, as the best allocation, the one that maximizes

$$\frac{\mathcal{L}_{\max}(\bar{\sigma}) - \mathcal{L}_{\max}(Pareto)}{\mathcal{C}(Pareto)}.$$

This solution looks for improved performance weighted by the cost of the reallocation.

4) Minimum Feasible: This strategy aims simply at finding a solution that does not overload any channel in the network. When there is no solution that avoids overloading, the strategy looks for the solution that minimizes the overload.

The first two simple strategies simply choose the solution that minimizes, independently one of the two criteria (cost and overload). Indeed, these two strategies propose as the best possible choice among the Pareto-optima solutions the solution of the two single-criterion problems respectively trying to minimize the maximum load and the reconfiguration cost.

IV. Results

We wish to analyze the different strategies previously proposed to assess the significance of the proposed approach. We focus on a toy network with $W = 4, N = 16$ since we are using CPLEX to obtain Pareto-optimal solutions; we further assume that the transmission speed on all wavelengths is the same, by convention normalized to 1.

To obtain the results, we randomly generate a traffic request matrix $\mathbf{T}$, where the element $t_{i,j}$, ranging from 0 to 1, represents the amount of traffic, normalized to wavelength capacity, node i is willing to transfer to node j. For each traffic request matrix randomly generated, we run CPLEX to obtain the Pareto-optimal solutions. Pareto-optimal solutions are locally efficient: none of the solutions is worse than other Pareto-optimal solutions in both cost and performance.

First, in Fig. 1 we report the number of Pareto-optimal solutions found by running the solution algorithm when generating 100 instances of the traffic matrix. The number of available solutions is fairly large, with an average number of roughly 6.5 Pareto-optimal solutions available in each instance, with a small number of situation in which up to 9 solutions are available. This is rather surprising if we consider the relatively small network size, and testifies that the problem we are studying is interesting to solve.

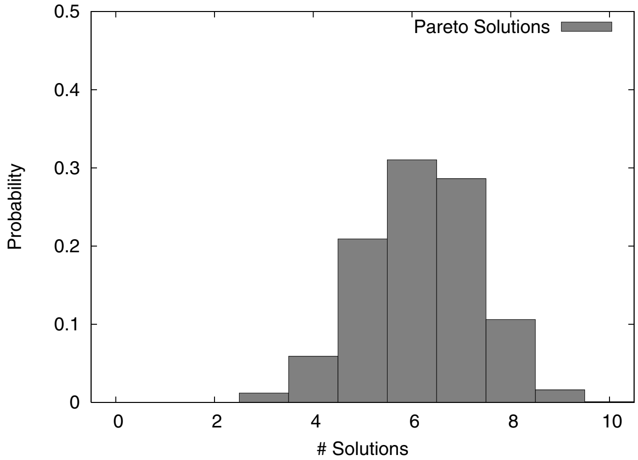

Fig. 1. Amount of Pareto Solutions : W = 4 N = 16

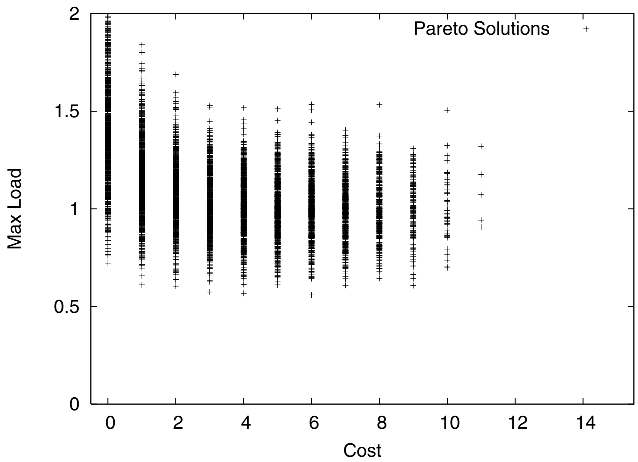

Fig. 2. Pareto Solutions : W = 4 N = 16

In Fig. 2 we show the characteristics of the Pareto-optimal solutions in terms of cost and performance. Again differences are quite astonishing: reconfiguration costs (measured in terms of number of receivers that must be tuned to ensure reallocation) range from 0 to 11, whereas the maximum load on a wavelength channel can reach overloading values close to 2.

Finally, Fig.3-6 show the characteristics of the solutions obtained when selecting the Pareto-optimal solutions according to the four previously described strategies. The Minimum Cost strategy shows how penalizing in terms of wavelength load can be the absence of any reconfiguration process. The Minimum Overload strategy, on the contrary, requires very high reconfiguration costs but it controls the maximum wavelength load very well, as expected. The Minimum Feasible strategy provides good values of maximum load, but the price in terms of reconfiguration cost is significant. The Best Ratio strategy seems to provide the best compromise, since loads are reasonably well kept under control with values comparable to those of Min Overload and Mean Feasible strategies, at reasonably small cost.

Clearly these results are only indicative of possible trends. The choice of the preferred strategy, as well as of the proper cost function, must be made by the network administrator taking into account the specific constraints of the network architecture and of the SLA offered to customers.

V. Conclusion

We have discussed why multi-criteria algorithms may play an important role in reconfiguration strategies on WDM networks. We have discussed the concept of Pareto-optimal solutions when looking at reconfiguration costs and performance as the two key issues when choosing whether a reconfiguration should take place or not.

We have introduced four possible strategies to guide the choice among Pareto-optimal solutions obtained by running CPLEX as an optimization tool, and we have discussed their performance.

We have also shown that the number of possible efficient Pareto-optimal solutions is fairly large, so that the problem is of practical significance.

The most promising reconfiguration strategy seems to be Best Ratio, which provides the best balance between reduced reconfiguration costs and controlled wavelength load.

References

[1] W. Golab, and R. Boutaba, “Policy-driven automated reconfiguration for performance management in WDM optical networks”, *IEEE Communication Magazine*, Vol. 42, No. 1, pp. 44-51, Jan. 2004

[2] I. Baldine, G. N. Rouskas, ”Dynamic load balancing in broadcast WDM networks with tuning latencies”, *Proc. IEEE Infocom*, Apr. 1998.

[3] M. Pinedo, “Scheduling : theory, algorithms, and systems”, *Prentice Hall*, 2002,

[4] R. L. Graham, “Bounds on multiprocessing timing anomalies”, *SIAM J. Appl. Math., vol 17, no 2, pp. 416-429, Mar. 1969*

[5] D. Vanderpooten, P. Vincke, ”Description and analysis of some representative interactive multicriteria procedures”, *Math. Comput, Model. 12*, pp 1121-1238, 1989.

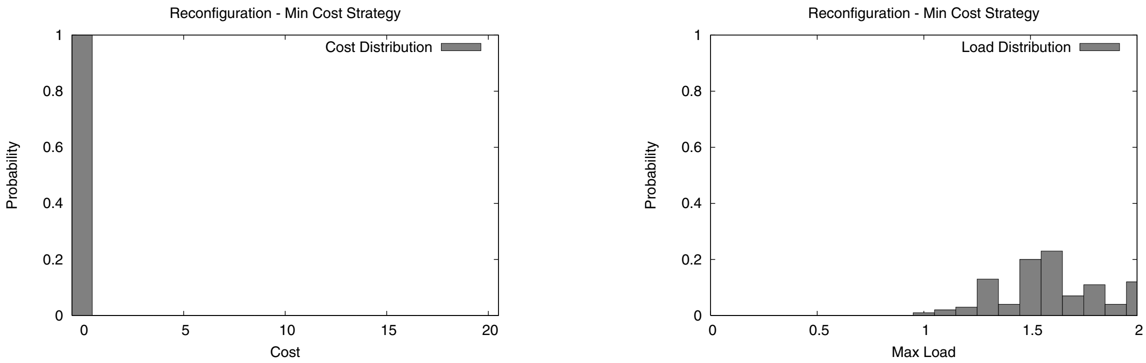

Fig. 3. Minimum Cost Strategy

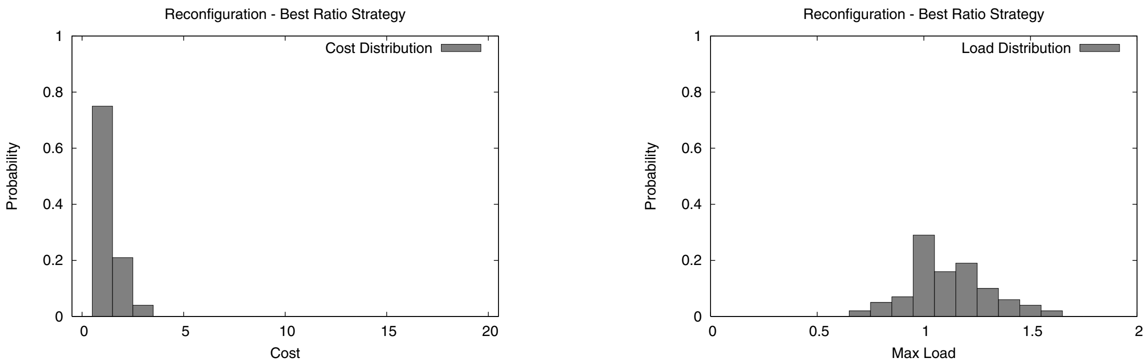

Fig. 4. Best Ratio Strategy

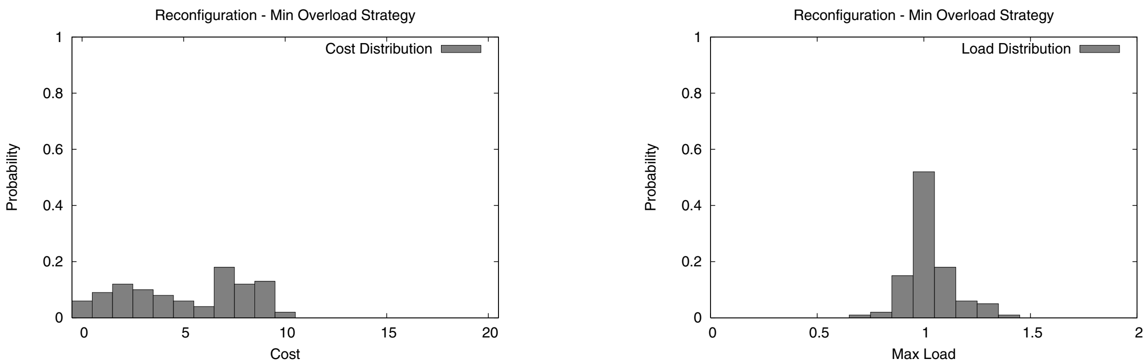

Fig. 5. Minimum Overload Strategy

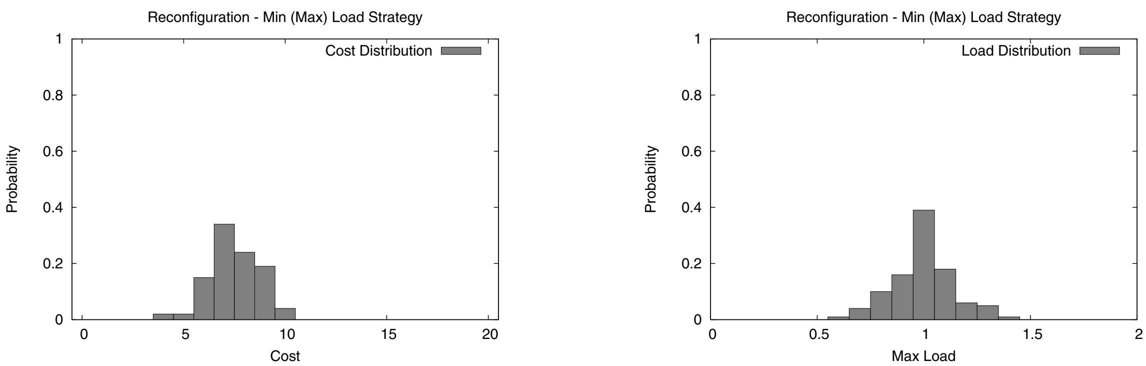

Fig. 6. Minimum Feasible Strategy